

# Water budget in the Baltic Sea drainage basin: Evaluation of simulated fluxes in a regional climate model

Petter Lind<sup>1)</sup> and Erik Kjellström<sup>1)2)\*</sup>

<sup>1)</sup> Swedish Meteorological and Hydrological Institute, SMHI, SE-601 76 Norrköping, Sweden  
(\*corresponding author's e-mail: Erik.Kjellstrom@smhi.se)

<sup>2)</sup> Department of Meteorology, Stockholm University, SE-106 91 Stockholm, Sweden

Received 18 Oct. 2007, accepted 9 Oct. 2008 (Editor in charge of this article: Harri Koivusalo)

Lind, P. & Kjellström, E. 2009: Water budget in the Baltic Sea drainage basin: Evaluation of simulated fluxes in a regional climate model. *Boreal Env. Res.* 14: 56–67.

We investigated the Rossby Centre regional climate model, RCA3, and its ability to reproduce the water budget of the Baltic Sea drainage basin during the period from 1979 to 2002. The model was forced on its lateral boundaries with European Centre for Medium-Range Weather Forecasts Re-Analysis data, ERA40. Simulated long-term means and inter-annual variability were compared with observational records and model-derived data. The basin-wide water fluxes were broadly captured by the model, and annual mean net precipitation over land agreed well (i.e., within 5%) with observed total discharge to the Baltic Sea. Long-term annual means of precipitation were around 20% higher in RCA3 compared with reference data, the differences being in most months statistically significant at the 5% level. On the other hand, differences between the reference datasets were evident and in most months also statistically significant. The inclusion of a high-resolution dataset showed a close agreement compared with RCA3; differences were less than 5% in the long-term annual mean. Therefore, more high-resolution observational datasets, especially for evaporation and runoff, are required to refine the water budget and compare water fluxes on sub-regional and local scales.

## Introduction

The highly heterogeneous terrain of the Baltic Sea drainage basin, including mountainous areas, vast coastlines, and a large semi-enclosed sea, presents a challenge and an opportunity in observing and modelling the water budget on a regional scale (Raschke *et al.* 2001). However, it is the regional and local characteristics of the precipitation and evaporation patterns in the Baltic Sea drainage area that make it difficult to produce accurate long-term records with sufficiently high spatio-temporal resolu-

tion. Moreover, undercatchment (low sampling) issues, especially in winter in the presence of snow and strong wind, further reduce the quality of the observations (Rubel and Hantel 2001). Atmosphere–ocean general circulation models (AOGCMs) with a relatively coarse resolution are unable to fully capture regional precipitation patterns having a strong seasonal cycle in this area (Graham *et al.* 2008). Dynamical downscaling of large-scale fields from AOGCMs by high-resolution regional climate models (RCMs) can provide such regional- and local-scale climate information (e.g., Christensen *et al.* 2007, Giorgi

*et al.* 2001). Provided that the large-scale circulation, temperature and humidity fields from the AOGCMs are realistic, RCMs can generate meaningful fine-scale structures that are absent in the AOGCM output (Christensen *et al.* 2007). Several RCMs have been used to investigate the spatial and temporal variability of the water budget of the Baltic Sea drainage area, for both shorter, relatively well-observed time periods (e.g., Jacob *et al.* 2001) and longer decadal and multi-decadal time series (e.g., Christensen *et al.* 1997, Jacob 2001, Ruprecht and Kahl 2003, Hagemann *et al.* 2004, Kjellström and Ruosteenoja 2007).

Previous RCM studies found that the models generally managed to reproduce the observed seasonal cycle in the Baltic Sea drainage basin to within the range of uncertainty given by observational datasets (Graham *et al.* 2008). Still, problems do exist with some model simulations, most notably, overestimated precipitation in winter in many simulations compared with observations (Kjellström and Ruosteenoja 2007). In previous studies, the boundary conditions were provided by either AOGCMs or reanalysis data. As the reanalysis data represent a combined product of results from a weather forecast model and observations, they follow the true state of the atmosphere in recent decades better than the AOGCMs do. Consequently, the problem of excess precipitation in winter may partly be attributed to systematic errors in the pressure fields of the driving AOGCMs (e.g., Kjellström and Ruosteenoja 2007). However, overestimation also occurs when using reanalysis data on the boundaries, implying that either errors in the reanalysis data or deficiencies in the model physics of the RCMs (Hagemann *et al.* 2004) may also play a role.

The earlier version of the reanalysis dataset used in previous RCM studies, ERA15 (Gibson *et al.* 1997), suffered from several problems. One of the most significant deficiencies in the hydrological cycle was the overestimated evaporation over land giving too negative net precipitation (i.e., precipitation minus evaporation). Furthermore, a dry bias was present in winter over Europe. The former problem has been significantly mitigated, and the latter almost entirely removed in the more recent ERA40 dataset

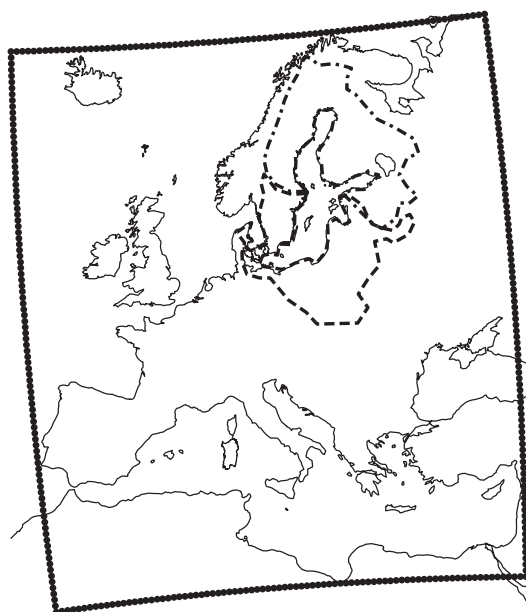
(Uppala *et al.* 2005). In addition, the snow-pack distribution is more realistic in ERA40 than in ERA15 and the cold wintertime bias over land in ERA15 is removed (Kållberg *et al.* 2005). Furthermore, water vapour in ERA40 validates well against independent observations of the total-column water vapour (Kållberg *et al.* 2005). In view of these improvements in the reanalysis data, evaluating the water budget in a regional climate model using ERA40 on the boundaries is justified.

This study evaluated the Rossby Centre regional climate model, RCA3 (Kjellström *et al.* 2005), forced by improved boundary conditions, and its ability to reproduce the atmospheric part of the water balance in the Baltic Sea drainage area. We examined how RCA3 quantified the water fluxes, i.e., precipitation, evaporation, and runoff, and the degree to which the observational information could be used to test this ability. This was done by comparing model results with available field measurements and with results from other models. By including several datasets regarding observed precipitation, we could assess model results in relation to the actual climate, while acknowledging some uncertainties in estimates based on measured data. In particular, we investigated how RCA3 reproduced monthly and annual fluxes and whether model results were within the natural variability evident in the observations.

## Materials and methods

### The RCA3 regional climate model

RCA3 is a regional climate model that originates from the high-resolution numerical weather prediction model HIRLAM (Källén 1996). RCA3 (Kjellström *et al.* 2005) builds on its predecessor, RCA2 (Jones *et al.* 2004), retaining its dynamic core but incorporating some major adjustments in the sub-grid-scale parameterization, for example, concerning convection and precipitation production. RCA3 includes a lake model, PROBE (Ljungemyr *et al.* 1996), and a new land-surface scheme (LSS) (Samuelsson *et al.* 2006). The LSS is constructed using a tile approach, including three tiles representing forest, open land, and



**Fig. 1.** The model domain over the Baltic Sea drainage basin in Europe. The basin is divided into the north-eastern (NE) part (dash-dotted line), and the south-western (SW) region (dashed line).

snow, the forest fraction being given by the MPI database (Hagemann *et al.* 1999). The forest tile is divided into three sub-tiles: forest canopy, forest floor soil with snow cover, and forest floor soil without snow cover. The open land tile is divided into vegetated and bare soil sub-tiles. In total, depending on the presence of snow, three to five surface energy balances are represented in each grid box. The surface fluxes of heat and momentum from the tiles are combined as a weighted grid average according to the fractional coverage of each tile in the grid box. The soil is divided into five layers with respect to tempera-

ture and is constrained by a no-flux boundary at the bottom.

## Model setup and reference data

In the experiment, RCA3 was run on a rotated latitude/longitude grid at a horizontal resolution of approximately 50 km ( $0.44^\circ \times 0.44^\circ$ ), with 24 hybrid levels up to 10 hPa in the vertical, and with an integration time step of 30 minutes. The model domain covered Europe (Fig. 1). Lateral boundary conditions, sea surface temperatures, and sea-ice conditions from ERA40 were updated in RCA3 every six hours. We investigated model results for 1979–2002, due to the relatively high density of observational datasets available for these years (Table 1). Annual, seasonal, and monthly fluxes of precipitation ( $P$ ), evaporation ( $E$ ), and runoff ( $R$ ) were compared with observations. Both RCA3 and observational data were area-integrated for all Baltic Sea drainage basin land points and for two separate parts of this region, namely, the northeastern (NE) and the southwestern (SW) parts (Fig. 1), following Graham *et al.* (2008).

Model results were compared with several observational datasets, ERA40, and data from a hydrological model, HBV (Table 1). The ERA40 and HBV datasets represent “quasi-observations” as they are derived partly from models. The simulated precipitation was compared with ERA40, CRU, SMHI, Rubel, and GPCP data. Global Precipitation Climatology Project (GPCP) data are based on a combination of gauge measurements and satellite information and constitute a global dataset. Climate Research Unit (CRU) data are based on gauge measurements (global coverage, land only). Swedish Meteorological and

**Table 1.** Observational data sets used in the analysis.

Data set	Resolution	Variables	Period	Source
ERA40 (quasi obs.)	$2^\circ \times 2^\circ$	$T_{2m}$ ; Precip.	1961–2002	Uppala <i>et al.</i> (2005)
GPCP vs. 2 (obs.)	$2.5^\circ \times 2.5^\circ$	Precipitation	1979–2005	e.g. Adler <i>et al.</i> (2003)
CRU TS 2.1 (obs.)	$0.5^\circ \times 0.5^\circ$	$T_{2m}$ ; Precip.	1901–2002	Mitchell and Jones (2005)
SMHI (obs.)	$1^\circ \times 1^\circ$	Precip. (SYNOP)	1980–2005	www.smhi.se/bhdc
HBV (quasi obs.)	$1^\circ \times 1^\circ$	Runoff; Evap.	1980–2005	Graham (1999)
Rubel data (obs.)	$1/6^\circ \times 1/6^\circ$	Bias corr. Precip.	1996–2000	Rubel and Hantel (2001)

Hydrological Institute (SMHI) data are based on SYNOP measurements. This dataset is gridded and covers the Baltic Sea drainage basin. The high-resolution Rubel data are based on over 4000 gauge measurements from the Baltic Sea drainage basin that were bias corrected using a precipitation and correction analysis model (Rubel and Hantel 2001). Even though the Rubel dataset covers a shorter period (1996–2000), its record may be considered most accurate and nevertheless valuable in the evaluation.

Evaluation of model performance in simulating evaporation is difficult due to its insufficient observational data coverage and high spatial variability. An alternative to the very sparse *in situ* measurements in the Baltic Sea drainage area is to compare simulated evaporation rates with results of hydrological models or reanalysis data. Here, evaporation was compared with outputs from the Swedish HBV Baltic Basin Water Balance Model (HBV) and ERA40. HBV is a hydrological model in which the physical processes are not implemented explicitly; instead, the model uses a conceptual approach in which the physical processes are statistically based simplifications suited to the area (Graham 1999).

Simulated runoff was also compared with outputs from HBV and ERA40 data. In HBV, runoff was calibrated against observed discharge, and the model was shown to be skilful in simulating integrated discharge to the Baltic Sea (Bergström *et al.* 2007). Therefore, HBV runoff may be used in evaluating RCA3 runoff. In ERA40, runoff is given by the net precipitation (i.e., precipitation minus evaporation,  $P - E$ ). In RCA3,  $R$  is calculated only for the land fractions of the grid points by the LSS.  $P$  and  $E$ , in contrast, are calculated for the entire grid box, including the land, lake, and sea fractions. Taken together, this means that the runoff in RCA3 is not identical to the simulated net precipitation over land.

Additionally, an observational record of total discharge into the Baltic Sea (Bergström and Carlsson 1994) was included in the evaluation. This record is not purely observational, as HBV results were used from 1997 to 2002 to estimate discharge in the Baltic Proper sub-catchment. No routing is included in RCA3, so integrated runoff from the Baltic Sea catchment cannot be directly compared with observed discharge into

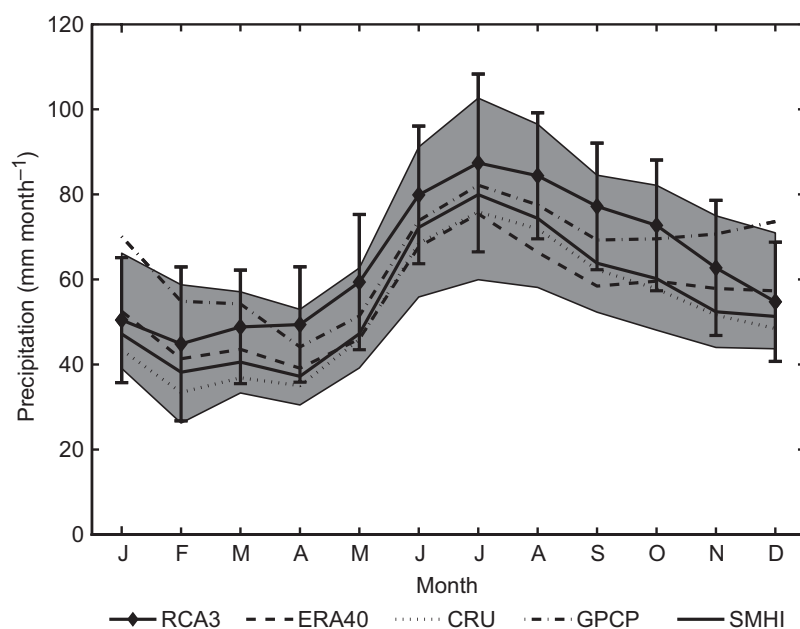
the Baltic Sea. Instead, we estimate total discharge to the Baltic Sea as net precipitation over the land area in the Baltic Sea drainage basin.

## Methods of comparison

We compared multi-year monthly and seasonal means of precipitation, evaporation, and runoff from the model with the corresponding reference values. Furthermore, the standard deviations of the simulated and reference data were calculated to estimate whether the inter-annual variability was captured by the model. Throughout the evaluation, *relative* not *absolute* differences between RCA3 and reference data are presented. In certain seasons with small absolute water fluxes, relative differences could still be quite large; this was particularly true for winter-time evaporation.

To quantify the uncertainties in the inferences from the data, we calculated confidence intervals using a bootstrap resampling method (Efron and Tibshirani 1993). After 10 000 random resamples of the data, two-tailed  $p$ -values for differences in means or standard deviations were calculated. The  $p$  values were used to test our null hypothesis: that the means (or standard deviations) of two components or variables, in this case the model and observation, are equal. We chose 0.05 to be the critical  $p$  value. Higher  $p$  values mean that we have to reject the null hypothesis at the 5% level, while lower  $p$  values imply that the simulated numbers do not differ from the reference data in a statistically significant way at the 5% level. Before the bootstrap, a test was performed to clarify whether or not the time series of monthly fluxes of simulated and reference data (observations and quasi-observations) were autocorrelated (i.e., serially correlated). The results indicated only limited autocorrelation, so no adjustment was made.

The test for statistical significance was not performed for comparisons with the Rubel data, as that dataset covers only a short time interval. In this case, we considered differences between the two datasets as being significant if the monthly mean of one dataset was outside the interval defined by the standard deviation of the other.



**Fig. 2.** Annual cycle of monthly mean precipitation averaged over the Baltic Sea drainage area for the years 1979–2002. The shaded area and the error bars represent inter-annual variability in SMHI and RCA3 data, respectively, in terms of  $\pm 1$  standard deviation (SD).

## Results

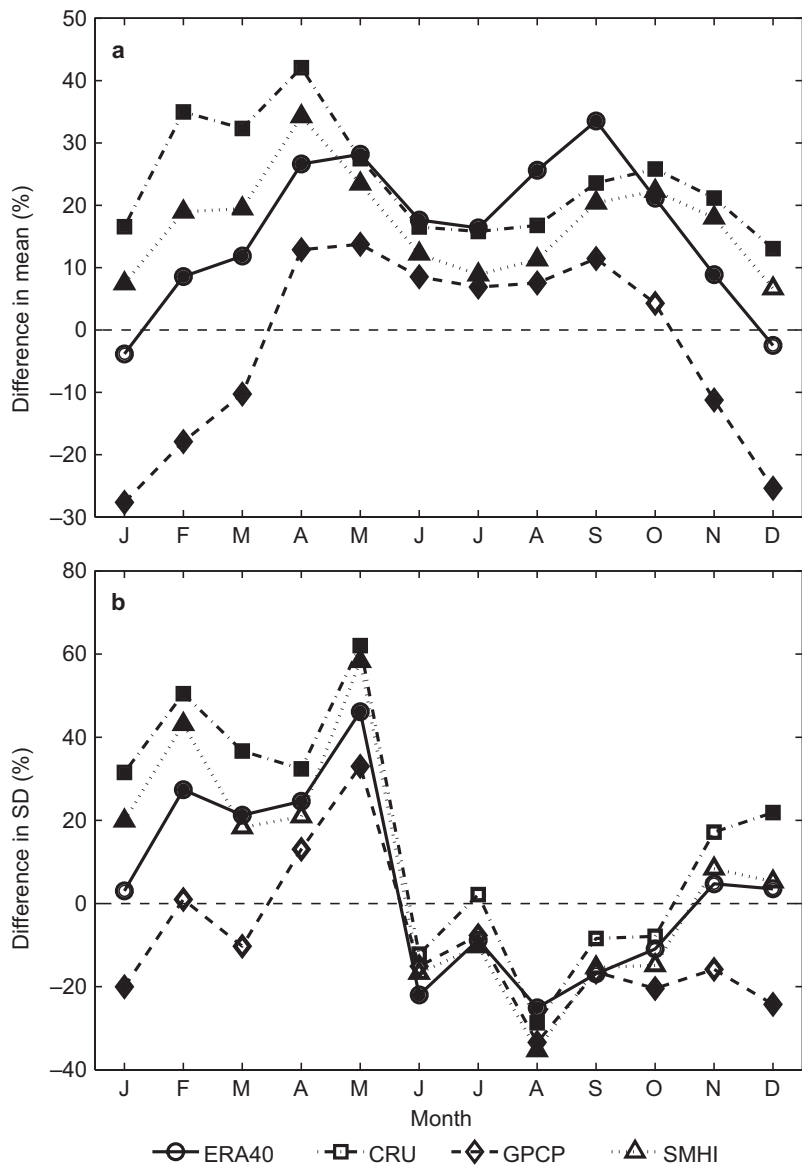
### Precipitation

The gross features of the seasonal cycle were well represented in the model, with maximum precipitation occurring in summer and minimum in February (Fig. 2 and Table 2). However, RCA3 generally simulated more precipitation over the year than almost all reference datasets. The relative deviations were less than  $\pm 25\%$  for most months, except in parts of spring and autumn when the overestimation was even larger (Fig. 3a). An exception was seen in winter, when RCA3 simulated significantly less precipitation than did GPCP. Differences between RCA3 and the reference datasets were statistically significant for all months, except December (in ERA40 and SMHI), January (ERA40), and October (GPCP). The long-term (1979–2002) annual mean was approximately 20% higher in RCA3 than in the reference datasets. For the shorter period (1996–2000), comparison with the high-resolution Rubel data indicated good agreement with RCA3, being only approximately 5% higher than the long-term annual mean.

Using the more detailed spatial information in the Rubel dataset, we looked at the ability of

**Table 2.** Seasonal and annual mean water fluxes as simulated in RCA3 (top row) for the period 1979–2002. The second row for each season shows the corresponding mean values of the reference data — numbers from left to right in the table refer to SMHI, ERA40 and CRU data for precipitation (*P*), and to HBV and ERA40 data for evaporation (*E*) and runoff (*R*).

	<i>P</i>	<i>E</i>	<i>R</i>
<b>Winter</b>			
Mean (mm month <sup>-1</sup> )	51	8	18
	46; 50; 42	2; 8	20; 25
Difference(%)	10; 2; 21	400; 0	-11; -39
<b>Spring</b>			
Mean (mm month <sup>-1</sup> )	53	37	44
	42; 43; 39	30; 32	49; 47
Difference(%)	27; 24; 36	23; 16	-11; -7
<b>Summer</b>			
Mean (mm month <sup>-1</sup> )	86	79	26
	75; 70; 72	68; 69	22; 19
Difference(%)	14; 22; 19	16; 14	18; 37
<b>Autumn</b>			
Mean (mm month <sup>-1</sup> )	71	33	22
	58; 59; 57	22; 25	19; 22
Difference(%)	22; 22; 25	50; 32	16; 0
<b>Annual</b>			
Mean (mm month <sup>-1</sup> )	65	39	27
	55; 56; 53	31; 34	28; 28
Difference(%)	19; 18; 23	26; 15	-4; -4



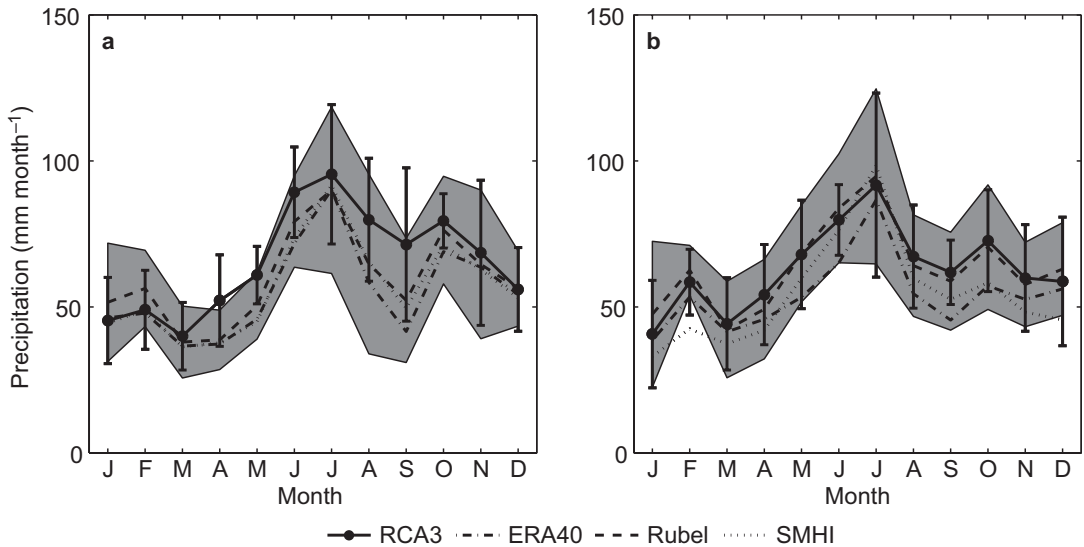
**Fig. 3.** Relative differences between RCA3 and reference data in (a) monthly mean precipitation, and (b) its standard deviation (SD) characterising inter-annual variability for the years 1979–2002. Filled markers represent differences that are statistically significant at the 5% level.

RCA3 to simulate regional details (Fig. 4). For reference, we also included the SMHI, ERA40, and CRU data. In this shorter period, the higher spring and autumn estimates in RCA3 relative to the reference data, primarily in April, May, and September, were again prominent in both the northeastern (NE) and southwestern (SW) parts of the Baltic Sea drainage area. In the SW, in contrast, RCA3 and the Rubel data were in good agreement in these seasons as well. Overall, in the SW, RCA3 precipitation differed not more than  $\pm 10\%$  from the Rubel data in nearly

all months. In the NE, RCA3 simulated approximately 10%–25% more precipitation than did the Rubel data; April and September stood out as the worst in this regard, as RCA3 gave over 25% more precipitation for these months. In April, RCA3 precipitation even exceeded the upper limit of the inter-annual variability as given by Rubel data for that month (Fig. 4). In winter, conversely, RCA3 simulated less precipitation than found in the Rubel dataset, approximately 15% less in both the SW and the NE.

Notably, in the NE, ERA40 and Rubel dif-





**Fig. 4.** Annual cycle of monthly mean precipitation in the (a) NE and (b) SW regions of the Baltic Sea drainage basin for the years 1996–2000. The shaded area and the error bars represent inter-annual variability in Rubel and RCA3, respectively, in terms of  $\pm 1$  standard deviation (SD) for each month.

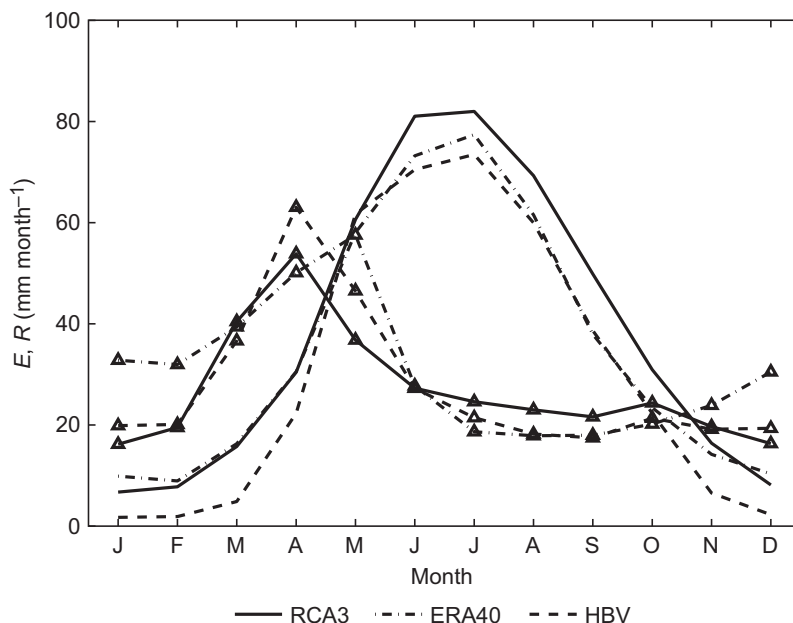
ferred no more than 15% from the SMHI data for all months. In contrast, RCA3 had 35%–45% more precipitation in spring and in September than did SMHI, the RCA3 monthly average being outside and above the interval defined by the SMHI inter-annual variability in April, May, and September (not shown). The difference was smaller in summer, being less than 25%. Therefore, in the NE, the spread between the reference datasets was less than the difference from RCA3, implying that RCA3 clearly overestimated precipitation in spring and, especially, in September.

The inter-annual variability was larger in RCA3 by 20%–60% than in most reference datasets in parts of the winter and spring (Fig. 3b). These differences were statistically significant. In summer and autumn, RCA3 tended to underestimate the inter-annual variability, although not to a statistically significant degree for most months. For the shorter period (1996–2000), the inter-annual variability in RCA3 was similar to that in Rubel, especially in the SW part. Compared with the other reference datasets, the inter-annual variability in Rubel was more similar to RCA3 in the SW, though this was not the case in the NE. In September, a local minimum in precipitation is observed in the NE according

both to Rubel and SMHI (Fig. 3b). The SMHI inter-annual variability was relatively large for this month; nevertheless, RCA3 exceeded the upper limit of the observed variability, indicating that the model did not capture the full strength of the local minimum in this month.

## Evaporation and runoff

Compared with HBV, RCA3 simulated 15%–50% more evaporation throughout the year except in the winter months (including March and November), when evaporation rates were approximately 2–4 times larger (Fig. 5 and Table 2). The agreement was generally better with ERA40: RCA3 values were within  $\pm 40\%$  ( $\pm 10\%$ ) of ERA40 values in winter and autumn (spring and summer). These differences in the mean values were statistically significant in all months except May for HBV and March and April for ERA40. Time series for evaporation for 1980–2002 indicated that the inter-annual variability in winter was higher in RCA3 than in the reference datasets (not shown). In summer, the differences in inter-annual variability between RCA3 and the reference datasets were smaller and not statistically significant.



**Fig. 5.** Annual cycles of monthly mean evaporation ( $E$ ) and runoff ( $R$ ; lines with triangles). Values are averaged over the Baltic Sea drainage area.

The phase of the seasonal runoff cycle as given by HBV was captured in RCA3, although the amplitude differed (Fig. 5). The spring peak was approximately 15% lower in RCA3, mainly due to the lower snow water equivalent in RCA3 (not shown). In ERA40, the spring peak in runoff occurred one month later (May) than in RCA3 or HBV, and again the spring peak in RCA3 was lower. RCA3 simulated approximately 25%–30% more runoff than did ERA40 from July to October, while the simulated runoff in winter was approximately 45% lower. For most months, the difference in the mean between RCA3 and the reference datasets was statistically significant. However, compensation of differences between seasons led to similar annual means in RCA3, ERA40, and HBV (Table 2). Furthermore, in all seasons there was an indication that the inter-annual variability of runoff was lower in RCA3 than in HBV, though it was larger than in ERA40 (not shown). These differences in inter-annual variability were generally not statistically significant.

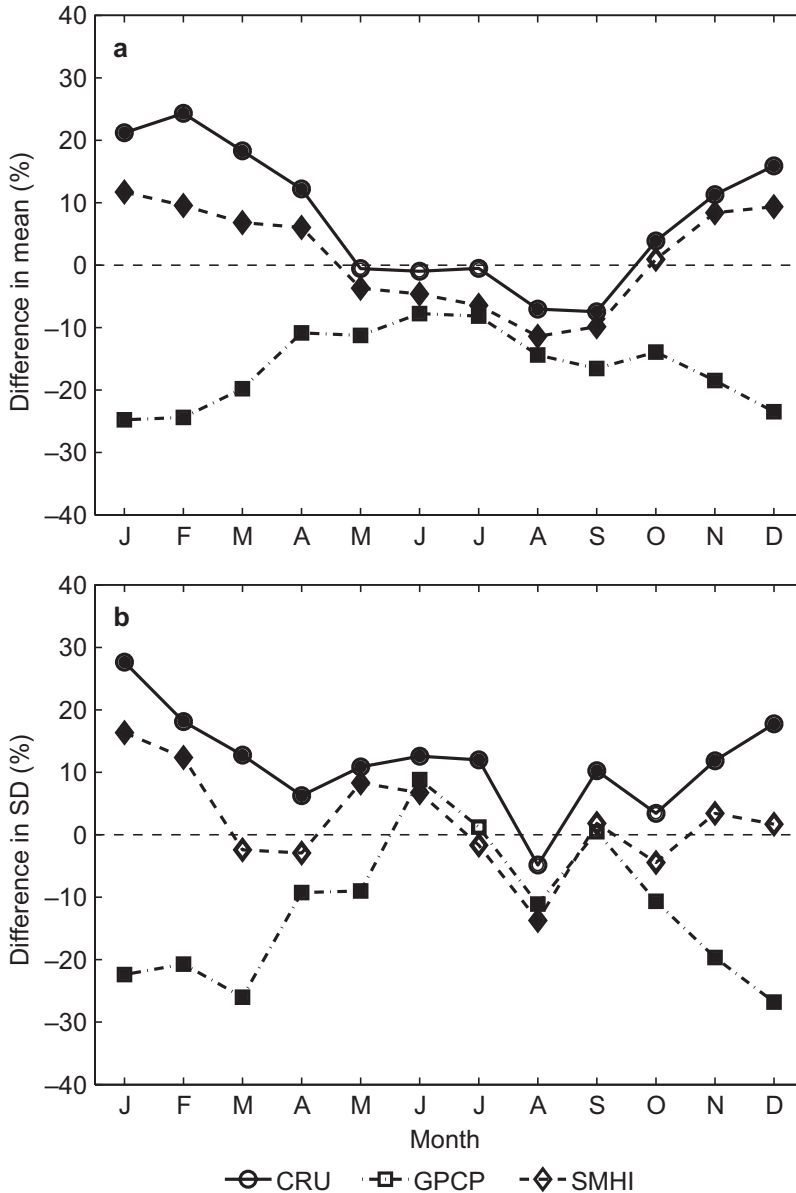
## Discussion

The statistical test showed that in nearly all months the differences in mean precipitation

between RCA3 and reference data in the Baltic Sea drainage basin were statistically significant. These results indicated a systematic wet bias in RCA3, as it simulated more precipitation than observed. Importantly, considerable differences were also evident between the various observational datasets, including ERA40. If these discrepancies are statistically significant, complications arise when drawing conclusions from the evaluation of RCA3 data in relation to the reference data. When ERA40 and GPCP data were used as references in two separate statistical tests, the simulated monthly means and the standard deviations characterising the inter-annual variability were found to differ significantly from the reference values in most months (Fig. 6). These differences between the reference datasets really emphasized the need to take into account the quality and possible deficiencies in the observational records when using these for model evaluation.

As mentioned earlier, the climate and topography of the Baltic Sea drainage basin region make it difficult to produce high-quality field measurements, especially in wintertime in the presence of snowfall and strong wind. This is also clear in Fig. 6, where the largest disagreements between the reference datasets are seen in winter and the smallest in summer. Throughout





**Fig. 6.** Relative differences between ERA40 and the other reference datasets in (a) monthly mean precipitation, and (b) its standard deviation (SD) characterising inter-annual variability for the years 1979–2002. Filled symbols represent differences that are statistically significant at the 5% level.

the year, the simulated precipitation in RCA3 was generally higher than the observed or ERA40 precipitation, primarily in spring and autumn in the NE part. An exception was the GPCP dataset, which indicated more precipitation in winter than did RCA3. The better agreement between RCA3 and the SMHI, the ERA40 and Rubel datasets suggest that GPCP is less reliable in this area in winter (Figs. 2 and 3a). This is supported by Rubel and Hantel (2001), who found that the bias correction coefficients introduced in the

GPCP dataset to deal with undercatchment were probably excessive leading to too overestimated precipitation in that dataset.

Another example of deficiencies in the quality of the observations or quasi-observations concerns ERA40; Hagemann *et al.* (2005) identified underestimated precipitation in the ERA40 dataset in the mid-latitudes. Their report stated that the too low precipitation estimates were related to spin-up effects, as they looked at the accumulated precipitation in the first six-hour

forecast (0–6). Therefore, here we used twelve-hour precipitation flux averages for the 12–24-hour forecast in ERA40.

The quality issue regarding observations was even more serious concerning evaporation and runoff. This was obvious, since no pure observational records could be used in the evaluation; only model-derived quasi-observations (ERA40 and HBV) provided data of sufficient spatio-temporal resolution. Furthermore, the HBV wintertime evaporation may be partially incorrect due to the inclusion of evaporation in a snow-fall correction factor (Graham and Jacob 2000). The reason for the larger evaporation in RCA3 compared with ERA40 in summer was unclear, and the difference may reflect differences in the surface parameterization schemes of the two models. In addition, as SMHI precipitation was used as input in HBV, larger evaporation rates in RCA3 were expected because simulated precipitation was higher than SMHI leading to more available water at the surface.

A clear bias in RCA3 precipitation has also been identified, and the missing September minimum was most prominent in the NE. This was not only evident in the shorter period (1996–2000), but also in the longer (1979–2002) SMHI and ERA40 time series (not shown). Also in this longer period, the September minimum was systematically missing in RCA3, likely due to the overestimated precipitation in summer. This led to wetter soils, more evaporation, and consequently more precipitation in late summer and early autumn. For the seasonal cycle of evaporation, RCA3 indicated higher values in summer and autumn than did HBV and ERA40 data, supporting this explanation.

Calculated multi-year annual means of integrated water fluxes over the Baltic Sea drainage area indicated that the net precipitation, i.e., precipitation minus evaporation, totalled  $512 \text{ km}^3 \text{ y}^{-1}$ . This was assumed to represent the total discharge into the Baltic Sea. At first sight comparison with the observed discharge indicated very good agreement; with an estimated total based on observations of  $517 \text{ km}^3 \text{ y}^{-1}$ , RCA3 differed less than 1% from the observations. But, this good agreement in annual mean conditions is a result of compensating seasonal errors in runoff (Fig. 5).

## Conclusions

The RCA3 water fluxes were broadly consistent with available observational and reanalysis datasets, both in the seasonal mean climate and the inter-annual variability. This implies that RCA3, when forced with boundary conditions from ERA40, could simulate the area-integrated fluxes of water over land areas in the Baltic Sea drainage basin. At a more detailed level however, the basin-wide long-term annual means of precipitation were generally overestimated in RCA3 by approximately 20% compared with most reference datasets. Looking at the seasonal cycle, differences in monthly mean precipitation between RCA3 and the reference data were statistically significant at the 5% level in most months. We also showed that the differences between the reference datasets themselves were mainly statistically significant. Therefore, we can not draw any strong conclusions regarding RCA3's ability to correctly simulate precipitation in the area. However, the close agreement between RCA3 and the high-resolution Rubel data, both on a monthly basis and in the long-term mean, lends confidence to the model results.

Evaluating the differences between RCA3 and quasi-observations regarding evaporation and runoff was more complicated, since the reference data were derived from other models. For example, differences in parameterizations may partly contribute to deviations in the results, so it was difficult to determine whether RCA3 or the quasi-observational datasets (ERA40 and/or HBV) were most realistic. Large seasonal differences between model and reference data sets can be hidden in the annual mean as shown for simulated integrated runoff and observed discharge into the Baltic Sea. A conclusion from this is that the fluxes in the water budget must be analysed on a seasonal basis when evaluating a model.

The inclusion of high-resolution observational records of precipitation was shown to provide valuable information on the differences in RCA3; however, the lack of high-quality, basin-wide information on evaporation and runoff generation posed serious difficulties in producing good model evaluations. In that regard, we conclude that better and more extensive gridded observational datasets for all hydrological vari-

ables in the water budget are needed in the future for these purposes. Such data may include radar-based precipitation observations and global positioning system-retrieved humidity profiles.

**Acknowledgements:** This research was partly funded by the European ENSEMBLES (GOCE-CT-2003-505539) project. The model simulation was made on the Tornado climate computer, funded by a grant from the Knut and Alice Wallenberg Foundation. The ERA40 data were provided by the European Centre for Medium range Weather Forecasts (ECMWF). We are grateful to Phil Graham who provided the river discharge data and Franz Rubel for providing the high-resolution precipitation data. Grigory Nikulin is acknowledged for performing the statistical tests.

## References

- Adler R.F., Huffman G.J., Chang A., Ferraro R., Xie P., Janowiak J., Rudolf B., Schneider U., Curtis S., Bolvin D., Gruber A., Susskind J. & Arkin P. 2003. The Version 2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeor.* 4: 1147–1167.
- Bergström S. & Carlsson B. 1994. River runoff to the Baltic Sea: 1950–1990. *Ambio* 23: 280–287.
- Bergström S., Jóhannesson T., Adalgeirsdóttir G., Ahlstrom A., Andreassen L.M., Andréasson J., Beldring S., Björnsson H., Carlsson B., Crochet P., de Woul M., Einarsson B., Elveheoy H., Flowers G.E., Graham L.P., Gröndal G.O., Gudmundsson S., Hellström S.-S., Hock R., Holmlund P., Jónsdóttir J.F., Pálsson F., Radic V., Reeh N., Roald L.A., Rogozova S., Rosberg J., Sigurdsson O., Suomalainen M., Thorsteinsson T., Vehviläinen B. & Veijalainen N. 2007. *Impacts of climate change on river runoff, glaciers and hydropower in the Nordic area*. Joint final report from the CE Hydrological Models and Snow and Ice Groups, CE-6, Hydrological Service, National Energy Authority, Reykjavík, Iceland.
- Christensen J.H., MACHENHAUER B., Jones R.G., Schär C., Ruti P.M., Castro M. & Visconti G. 1997. Validation of present-day regional climate simulations over Europe: LAM simulations with observed boundary conditions. *Clim. Dyn.* 13: 489–506.
- Christensen J.H., Hewitson B., Busiuc A., Chen A., Gao X., Held I., Jones R., Kolli R.K., Kwon W.-T., Laprise R., Magaña Rueda V., Mearns L., Menéndez C.G., Räisänen J., Rinke A., Sarr A. & Whetton P. 2007. Regional Climate Projections. In: Solomon S., Qin D., Manning M., Chen Z., Marquis M., Averyt K.B., Tignor M. & Miller H.L. (eds.), *Climate change 2007: the physical science basis*, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 847–940.
- Efron B. & Tibshirani R.J. 1993. *An introduction to the bootstrap*. Chapman and Hall, New York.
- Giorgi F., Hewitson B., Christensen J., Hulme M., Von Storch H., Whetton P., Jones R., Mearns L. & Fu C. 2001. Regional climate information — evaluation and projections. In: Houghton J.T., Ding Y., Griggs D.J., Noguer M., van der Linden P.J., Dai K., Maskell K. & Johnson C.A. (eds.), *Climate change 2001: the scientific basis*, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, U.K., pp. 583–638.
- Graham L.P., Chen D., Christensen O.B., Kjellström E., Krysanova V., Meier H.E.M., Radziejewski M., Rockel B., Ruosteenoja K. & Räisänen J. 2008. Projections of future climate change. In: The BACC Author Team (eds.), *Assessment of climate change for the Baltic Sea basin*, Springer, Berlin, pp. 133–219.
- Graham P. & Jacob D. 2000. Using large-scale hydrologic modeling to review runoff generation processes in GCM climate models. *Meteorol. Z.* 9: 49–57.
- Graham P. 1999. Modeling runoff to the Baltic Sea. *Ambio* 28: 328–334.
- Gibson J.K., Källberg P., Uppala S., Hernandez A., Nomura A. & Serrano E. 1997. *Era description*. ECMWF Reanal, Proj. Rep. Ser. 1, European Centre for Medium-Range Weather Forecasts, Reading, U.K.
- Hagemann S., Botzet M., Dümenil L. & Machenhauer B. 1999. *Derivation of global GCM boundary conditions from 1 km land use satellite data*. Tech. Rep. 289, Max-Planck-Institute for Meteorology, Hamburg, Germany.
- Hagemann S., Machenhauer B., Jones R., Christensen O.B., Déqué M., Jacob D. & Vidale P.L. 2004. Evaluation of water and energy budgets in regional climate models applied over Europe. *Clim. Dyn.* 23: 547–567.
- Hagemann S., Arpe K. & Bengtsson L. 2005. *Validation of the hydrological cycle of ERA-40*. ERA-40 Project Report Series No. 24, European Centre for Medium-Range Weather Forecasts, Reading, U.K.
- Jacob D., Van den Hurk B.J.J.M., Andrae U., Elgered G., Fortelius C., Graham L.P., Jackson S.D., Karstens U., Köpken Chr., Lindau R., Podzun R., Rockel B., Rubel F., Sass B.H., Smith R.N.B. & Yang X. 2001. A comprehensive model inter-comparison study investigating the water budget during the BALTEX-PIDCAP period. *Meteorol. Atmos. Phys.* 77: 19–43.
- Jacob D. 2001. A note to the simulation of the annual and inter-annual variability of the water budget over the Baltic Sea drainage basin. *Meteorol. Atmos. Phys.* 77: 61–73.
- Jones C.G., Ullerstig A., Willén U. & Hansson U. 2004. The Rossby Centre regional atmospheric climate model (RCA). Part I: Model climatology and performance characteristics for present climate over Europe. *Ambio* 33: 199–210.
- Kjellström E. & Ruosteenoja K. 2007. Present-day and future precipitation in the Baltic Sea region as simulated in a suite of regional climate models. *Clim. Change* 81: 28–291.
- Kjellström E., Bärring L., Gollvik S., Hansson U., Jones C., Samuelsson P., Rummukainen M., Ullerstig A., Willén U. & Wyser K. 2005. *A 140-year simulation of Euro-*

- pean climate with the new version of the Rossby Centre regional atmospheric climate model (RCA3). SMHI Reports Meteorology and Climatology No. 108, SMHI, Norrköping, Sweden.
- Kållberg P., Berrisford P., Hoskins B., Simmons A., Uppala S., Lamy-Thépaut S. & Hine R. 2005. *ERA-40 Atlas*. ERA-40 Project Report Series 19, European Centre for Medium Range Weather Forecasts, Reading, U.K.
- Källén E. 1996. *HIRLAM documentation manual: system 2.5*. Technical Report HIRLAM, SMHI, Norrköping, Sweden.
- Ljungemyr P., Gustafsson N. & Omstedt A. 1996. Parameterization of lake thermodynamics in a high-resolution weather forecasting model. *Tellus* 48A: 608–621.
- Mitchell T.D. & Jones P.D. 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int. J. Climatol.* 25: 69–712.
- Raschke E., Meywerk J., Warrach K., Andrae U., Bergström S., Beyrich F., Bosveld F., Bumke K., Fortelius C., Graham L.P., Gryning S.-E., Halldin S., Hasse L., Heikinheimo M., Isemer H.-J., Jacob D., Jauja I., Karlsson K.-G., Keevalik S., Koistinen J., van Lammeren A., Lass U., Launianen J., Lehmann A., Liljebladh B., Lobmeyr M., Matthäus W., Mengelkamp T., Michelson O.B., Napiórkowski J., Omstedt A., Piechura J., Rockel B., Rubel F., Ruprecht E., Smedman A.-S. & Stigebrandt A. 2001. The Baltic Sea Experiment (BALTEX): a European contribution to the investigation of the energy and water cycle over a large drainage basin. *Bull. Am. Meteorol. Soc.* 82: 2389–2413.
- Rubel F. & Hantel M. 2001. BALTEX 1/6-degree daily precipitation climatology. *Meteorol. Atmos. Phys.* 77: 155–166.
- Ruprecht E. & Kahl T. 2003. Investigation of the atmospheric water budget of the BALTEX area using NCEP/NCAR reanalysis data. *Tellus* 55A: 426–437.
- Samuelsson P., Gollvik S. & Ullerstig A. 2006. *The land-surface scheme of the Rossby Centre regional atmospheric climate model (RCA3)*. Report in Meteorology 122, SMHI, Norrköping, Sweden.
- Uppala S.M., Kållberg P.W., Simmons A.J., Andrae U., da Costa Bechtold V., Fiorino M., Gibson J.K., Haseler J., Hernandez A., Kelly G.A., Li X., Onogi K., Saarinen S., Sokka N., Allan R.P., Andersson E., Arpe K., Balmaseda M.A., Beljaars A.C.M., van de Berg L., Bidlot J., Bormann N., Caires S., Chevallier F., Dethof A., Dragosavac M., Fisher M., Fuentes M., Hagemann S., Holm E., Hoskins B.J., Isaksen I., Janssen P.A.E.M., Jenne R., McNally A.P., Mahfouf J.-F., Morcrette J.-J., Rayner N.A., Saunders R.W., Simon P., Sterl A., Trenberth K.E., Untch A., Vasiljevic D., Viterbo P. & Woollen J. 2005. The ERA-40 reanalysis. *Q. J. Roy. Meteorol. Soc.* 131: 2961–3012.